

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 16:09:05 ; Search time 299.73 seconds
(without alignments)
24.636 Million cell updates/sec

Title: US-09-331-631a-1_COPY_186_248

Perfect score: 355
Sequence: 1 KRDPQOREYEDCRRCRCEQGE.....MNPORGSGRYEEGEEES 63

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :
1: SPREMBL_15:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	355	100.0	666	10 Q9SP15	Q9SP15 macadamia l
2	332	93.5	666	10 Q9SP14	Q9SP14 macadamia l
3	320	90.1	625	10 Q9SP13	Q9SP13 macadamia l
4	125.5	35.4	593	10 Q9SEW4	Q9SEW4 juglans reg
5	120	33.8	525	10 Q43358	Q43358 theobroma c
6	112	31.5	637	10 Q03678	Q03678 hordeum vul
7	92	25.9	810	10 Q9ZM13	Q9ZM13 cucurbita m
8	87	24.5	539	4 Q9N0A2	Q9N0A2 homo sapien
9	86	24.2	1655	5 Q24754	Q24754 drosophila
10	82	23.1	1312	5 Q9VJ08	Q9VJ08 drosophila
11	82	23.1	1323	5 Q9VJ09	Q9VJ09 drosophila
12	81.5	23.0	242	10 Q81260	Q81260 tripsacum d
13	81.5	23.0	246	10 Q81261	Q81261 tripsacum d
14	80.5	22.7	296	10 Q9S709	Q9S709 arabidopsis
15	79.5	22.4	236	10 Q81254	Q81254 zea mays su
16	79	22.3	411	5 P91419	P91419 caenorhabdi
17	79	22.3	630	5 Q9W4J3	Q9W4J3 drosophila
18	79	22.3	1295	4 Q9P2R8	Q9P2R8 homo sapien
19	79	22.3	1300	11 Q9Jm92	Q9Jm92 mus musculu

20	79	22.3	1303	4 Q9PIX1	Q9PIX1 homo sapien
21	79	22.3	1308	11 Q9Jm52	Q9Jm52 mus musc
22	78.5	22.1	1128	5 Q9V260	Q9V260 drosophila
23	77.5	21.8	1483	5 Q9VD46	Q9VD46 drosophila
24	77	21.7	153	5 Q9VPM6	Q9VPM6 drosophila
25	77	21.7	646	2 Q9KM25	Q9KM25 vibrio chol
26	76.5	21.5	1089	12 Q40947	Q40947 kaposi's sa
27	76	21.4	242	10 Q81251	Q81251 zea mays su
28	76	21.4	242	10 Q9SBE6	Q9SBE6 zea mays su
29	76	21.4	720	5 Q23847	Q23847 dictyosteli
30	76	21.4	906	4 Q9UN21	Q9UN21 homo sapien
31	76	21.4	2123	5 Q9U957	Q9U957 dictyosteli
32	75.5	21.3	3469	5 Q9U412	Q9U412 drosophila
33	75.5	21.3	330	5 Q81118	Q81118 caenorhabdi
34	75.5	21.3	535	5 Q9U1N0	Q9U1N0 chironomus
35	75.5	21.3	556	5 Q76940	Q76940 drosophila
36	75.5	21.3	663	5 Q9W352	Q9W352 drosophila
37	75	21.1	238	10 Q81257	Q81257 zea luxuria
38	75	21.1	393	10 Q9ZTP0	Q9ZTP0 oryza sativ
39	75	21.1	1670	5 Q23901	Q23901 dictyosteli
40	74	20.8	238	10 Q81258	Q81258 zea luxuria
41	74	20.8	238	10 Q9SBE5	Q9SBE5 zea luxuria
42	74	20.8	238	10 Q9SBE6	Q9SBE6 zea luxuria
43	74	20.8	923	4 Q60424	Q60424 homo sapien
44	74	20.8	1038	5 Q60983	Q60983 dictyosteli
45	74	20.8	1129	12 Q9QR71	Q9QR71 kaposi's sa

ALIGNMENTS

RESULT 1

Q9SP15 PRELIMINARY; PRT; 666 AA.
ID Q9SP15
AC Q9SP15;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE VICILIN PRECURSOR.
GN AMP2.
OS Macadamia integrifolia (Macadamia nut).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; Proteaceae; Macadamia.
OX NCBI_Taxid=60698;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=NOT KNOWN.
RA Marcus J.P., Goulter K.C., Green J.L., Manners J.M.;
RT "A family of antimicrobial peptides is produced by processing of a 7S
RT globulin protein in Macadamia integrifolia kernels.";
RL Plant J. 0:0-(1999).
DR EMBL: AF161883; AAD54244.1; -
DR HSSP: P02853; 2PHL.
DR INTERPRO: IPR001113; -
DR PRAM: PF00546; Seedstore_7s; 1.
SQ SEQUENCE 666 AA; 78217 MW; C752B84BDFO224 CRC64;

Query Match 100.0%; Score 355; DB 10; Length 666;
Best Local Similarity 100.0%; Pred. No. 4.7e-35;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRDPQOREYEDCRRCRCEQGE.....MNPORGSGRYEEGEE 60
|||
Db 186 KRDPQOREYEDCRRCRCEQGE.....MNPORGSGRYEEGEE 245

QY 61 EQS 63
|||
Db 246 EQS 248

RESULT 2
Q9SP14

ID	Q9SP14	PRELIMINARY;	PRT;	666 AA.
AC	Q9SP14			
DT	01-MAY-2000 (TREMBLrel_13, Created)			
DT	01-MAY-2000 (TREMBLrel_13, Last sequence update)			
DT	01-OCT-2000 (TREMBLrel_15, Last annotation update)			
DE	VICILIN PRECURSOR.			
GN	AMP2.			
OS	Macadamia integrifolia (Macadamia nut).			
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;			
OC	Manoliophyta; eudicotyledons; Proteaceae; Macadamia.			
OX	NCBI_TaxID=60698;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=NOT KERNEL;			
RA	Marcus J.P., Goulter K.C., Green J.L., Manners J.M.;			
RT	"A family of antimicrobial peptides is produced by processing of a 7S			
RT	globulin protein in Macadamia integrifolia.";			
RL	Plant J. 0:0-0(1999).			
DR	EMBL; AF161884; AAD54245.1; -.			
DR	HSSP; P02853; 2PBL.			
DR	INTERPRO: IPR001113; -			
DR	PFAM; PF00546; Seedstore_7s; 1.			
SQ	SEQUENCE 666 AA: 78243 MW; 0ECN22E8710F8A7B CRC64;			

Query Match	93.5%	Score 332;	DB 10;	Length 666;
Best Local Similarity	92.1%;	Pred. NO. 2.8e-32;		
Matches 58; Conservative	4;	Mismatches 1;	Indels 0;	Gaps 0

```

QY 1 KRDPQQRREYEDCRRRCRQEQEPERQHQCLRCRQEQRQHGRCGDMNPNPQRCSGRYEGEE 60
      |||||
DB 186 KRDPQQRREYEDCRRRCRQEQEPERQHQCLRCRQEQRQHGRCGDLINPQRCSGRYEGEE 245
      |||||

```

QY	61	EQS	63
Db	246	KQS	248

SEQUENCE	FT	Q
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100
101	101	101
102	102	102
103	103	103
104	104	104
105	105	105
106	106	106
107	107	107
108	108	108
109	109	109
110	110	110
111	111	111
112	112	112
113	113	113
114	114	114
115	115	115
116	116	116
117	117	117
118	118	118
119	119	119
120	120	120
121	121	121
122	122	122
123	123	123
124	124	124
125	125	125
126	126	126
127	127	127
128	128	128
129	129	129
130	130	130
131	131	131
132	132	132
133	133	133
134	134	134
135	135	135
136	136	136
137	137	137
138	138	138
139	139	139
140	140	140
141	141	141
142	142	142
143	143	143
144	144	144
145	145	145
146	146	146

Query Match	90.1%;	Score 320;	DB 10;	length 625;
Best Local Similarity	88.9%;	Pred. No. 7.5e-31;		
Matches 56;	Conservative 4;	Mismatches 3;	Indels 0;	Gaps 0

OY 1 KRPQQREFEDCRKRCEQQEPHQHCLCKRCEQRQHSGGGDMNPPQRGSGRYEEGEE 60
 ||||| : |||||
D6 145 KRDPQQRFEVDCRRHCQEQLLYCGRRCQEQQRQHGRGDLMNPQRGSGRYEAGEE 204

QY	61	EQS	63
		:11	
Db	205	KQS	207

RESULT	4
Q9SEW4	
ID	Q9SEW4
AC	PRELIMINARY;
DT	01-MAY-2000 (TREMBLrel. 13, Created)
DT	01-MAY-2000 (TREMBLrel. 13, last sequence update)
DT	01-OCT-2000 (TREMBLrel. 15, last annotation update)
DE	VICILIN-LIKE PROTEIN PRECURSOR (FRAGMENT).
OS	Juglans regia (English walnut).
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC	Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC	Fagales; Juglandaceae; Juglans.
OX	NCBI_TaxID=51240;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	STRAIN=CV. SUNLAND; TISSUE=SOMATIC EMBRYO LINE;
RC	Teuber S.S., Jarvis K.C., Peterson W.R., Dandekar A.M., Ansari A.A.;
RT	"Identification and cloning of a cDNA encoding a vicilin-like protein
RT	Jug r 2, from English walnut kernel (Juglans regia): a major food
RT	allergen.";
RL	Submitted (MAY-1998) to the EMBL/Genbank/DBJ databases.
DR	EMBL; AF066055; AAF18269.1; -.
DR	HSSP; P02853; 2PHT.
DR	INTERPRO; IPR001113; -.
DR	PFAM; PF00546; Seedstore_7s; 1.
FT	NON_TER
FT	SEQUENCE
Q9	593 AA: 69990 MW: 9BA127E19B18C0D8 CRC64:

Query Match	35.48;	Score 125.5;	DB 10;	Length 593;
Best Local Similarity	45.98;	Pred. No. 2.2e-07;		

```

0y 1 KRDPQOREVEDCRRR--EODEPRQOHQCLRC-REORQHR-GDDMMNPORGSGAYE 56
   ::|||:|:|||::|:|||::|:|||::|:|||::|:|||::|:|||::|:|||
Db 116 RQDPQO-QYHRCQRRCIQEOSPERRQCCQRCREYKKEQGGREPREASPRRESRGE 174

```

QY	57	E	57
		—	
Db	175	E	175

RESULT	5		
043358			
ID	043358	PRELIMINARY;	PRT; 525 AA.
AC	043358;		
DT	01-NOV-1996 (Tremblrel, 01, Created)		
DT	01-NOV-1996 (Tremblrel, 01, Last sequence update)		
DT	01-OCT-2000 (Tremblrel, 15, Last annotation update)		
DE	VICILIN PRECURSOR.		
GN	CSV.		
OS	Theobroma cacao (Cacao).		
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;		
OC	Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;		
OC	Malvales; Malvaceae; Theobroma.		
OX	NCBI_TaxID=3641;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RT	TISSUE-LEAVES;		
RX	MEDLINE=92288309; PubMed=1600151;		
RA	Chenmy L., Fritz P.D.;		
RT	"Comparison of the structure and nucleotide sequences of vicilin genes		
RT	of cocoa and cotton raise questions about vicilin evolution.";		
RL	Plant Mol. Biol. 18:1173-1176(1992).		
DR	EMBL; X62625; CAA44493.1; -		
DR	EMBL; X62626; CAA44494.1; -		
HSSP	P02853; ZPHL.		

RA Brandon R.C., Rogers V.-H.C., Blazewicz R.G., Champe M., Pfeiffer B.D.,
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 April J.F., Agdayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 Baillieu R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolchakov S.,
 Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 Rosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 Hostin D., Houston K., Howland T.J., Wei M.-H., Idegawa C.,
 Jajalil M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Kelchum K.A.,
 Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
 Mount S.M., Moy M., Murphy B., Murphy L., Munz D.M., Nelson D.L.,
 Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacht J.M.,
 Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 Spier E., Spradling K.A., Stapleton M., Strong R., Sun E.,
 Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 "The genome sequence of *Drosophila melanogaster*.";
 Science 287:2185-2195(2000).
 DR EMBL: AEO03652; AAF53546.1; -
 DR FLYBASE: FBgn0032592; CG4668.
 DR INTERPRO: IPR001777; -
 DR INTERPRO: IPR001865; -
 DR PFAM: PF00041; fn3; 8.
 DR PRINTS: PRO0014; ENTPEIT1.
 DR PROSITE: PS00962; RIBOSOMAL_S2.1; UNKNOWN_1.
 DR SEQUENCE 1323 AA; 144904 MW; 937642564f61532D CRC64;

Query Match 23.1%; Score 82; DB 5; Length 1323;
 Best Local Similarity 36.2%; Pred. No. 0.082;
 Matches 21; Conservative 6; Mismatches 19; Indels 12; Gaps 2;

OY 14 RRCRCEQERQOHCQRCRQOQOGRGQDMNPRGSGS-----GREGESEERGS 63
 DB 217 RRLPQOHGOAHQ----VAQOQOQOARNGNPQHPRAGSSVGVSGTSEDEDMSS 270

RESULT 12
 ID 081260 PRELIMINARY; PRT: 242 AA.
 AC 081260;
 DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE GLOBULIN-1 (FRAGMENT).
 OS Tripsacum dactyloides (Gama grass).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; Liliopsida; Poales; Poaceae; Tripsacum.
 OC NCBI_TaxID=4563;
 OX 11
 RN SEQUENCE FROM N.A.
 RP Hilton H., Gaut B.S.;
 RT "Speciation and domestication in maize and its wild relatives:
 RT evidence from the Globulin-1 gene.";
 RL Genetics 0:0-0(1998).
 DR EMBL: AF064234; AAC31477.1; -
 DR HSP; P50477; ICAU.

DR MENDEL: 31902; Trida;1188;31902.
 DR INTERPRO: IPR001113; -
 DR PFAM: PF00546; Seedstore_7s; 1.
 FT NON_TER 242 242
 SQ SEQUENCE 242 AA; 27810 MW; CF1A3BFA73429776 CRC64;

Query Match 23.0%; Score 81.5; DB 10; Length 242;
 Best Local Similarity 35.2%; Pred. No. 0.02;
 Matches 19; Conservative 8; Mismatches 24; Indels 3; Gaps 1;

OY 12 CRRRCRCEQERQOHCQRCRQOQOGRGQDMNPRGSGS-----GSGRYEGEERQ 62
 DB 38 CARRCEDRPMHQPRCLQECREEREXQOERSRHEADRREGSGSGDERERERQ 91

RESULT 13
 ID 081261 PRELIMINARY; PRT: 246 AA.
 AC 081261;
 DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE GLOBULIN-1 (FRAGMENT).
 OS Tripsacum dactyloides (Gama grass).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; Liliopsida; Poales; Poaceae; Tripsacum.
 OC NCBI_TaxID=4563;
 RN 11
 RP SEQUENCE FROM N.A.
 RP Hilton H., Gaut B.S.;
 RT "Speciation and domestication in maize and its wild relatives:
 RT evidence from the Globulin-1 gene.";
 RL Genetics 0:0-0(1998).
 DR EMBL: AF064235; AAC31478.1; -
 DR HSP; P50477; ICAU.
 DR MENDEL: 31903; Trida;1188;31903.
 DR INTERPRO: IPR000901; -
 DR INTERPRO: IPR001113; -
 DR PFAM: PF00546; Seedstore_7s; 1.
 DR PROSITE: PS00667; CPASASE_2; UNKNOWN_1.
 FT NON_TER 246 246
 SQ SEQUENCE 246 AA; 28408 MW; 8D9B45991P9C5FB9 CRC64;

Query Match 23.0%; Score 81.5; DB 10; Length 246;
 Best Local Similarity 35.2%; Pred. No. 0.02;
 Matches 19; Conservative 8; Mismatches 24; Indels 3; Gaps 1;

OY 12 CRRRCRCEQERQOHCQRCRQOQOGRGQDMNPRGSGS-----GSGRYEGEERQ 62
 DB 38 CARRCEDRPMHQPRCLQECREEREXQOERSRHEADRREGSGSGDERERERQ 91

RESULT 14
 ID 09S709 PRELIMINARY; PRT: 296 AA.
 AC 09S709;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
 DE T22C5.10.
 GN T17H3.14.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OC NCBI_TaxID=3702;
 OX 11
 RN SEQUENCE FROM N.A.
 RP Chao Q., Brooks S., Buehler E., Johnson-Hopson C., Khan S., Kim C.,
 RA Shinn P., Altafi H., Bel Q., Chin C., Chiu J., Choi E., Conn L.,
 RA Conway A., Gonzales A., Hansen N., Howng B., Koo T., Lam B., Lee J.,

RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharly N., Nguyen M.,
 RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaverl A.,
 RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
 RA Ecker J.R.,
 RT "Genomic sequence for Arabidopsis thaliana BAC T22C5 from chromosome
 RT I.";
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Vysotskaia V.S., Schwartz J.R., Yu G., Toriumi M., Lenz C., Liu S.,
 RA Lee J.M., Liu A., Li J., Gonzalez A., Liu K., Vaysberg M., Sakano H.,
 RA Chin C., Howng B., Choi E., Chlou J., Altafi H., Araujo R., Brooks S.,
 RA Buehler E., Chao Q., Conn L., Conway A.B., Dunn P., Hansen N.,
 RA Huizar L., Khan S., Kim C., Palm C., Rowley D., Shinn P., Walker M.,
 RA Davis R.W., Ecker J.R., Federspiel N.A., Theologis A.,
 RT "Arabidopsis thaliana chromosome 1 BAC T17H3 sequence.";
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC012375; AAF24943.1; -;
 DR EMBL: AC005916; AAD46002.1; -;
 DR INTERPRO: IPR000504; -;
 DR INTERPRO: IPR000571; -;
 DR PFAM: PF00076; ttm; 1.
 DR PFAM: PF00642; zf-CCCH; 2.
 SQ SEQUENCE 296 AA; 34572 MW; F284D29BF68445BF CRC64;

Query Match 22.7%; Score 80.5; DB 10; Length 296;

Best Local Similarity 34.9%; Pred. No. 0.032; Mismatches 33; Indels 1; Gaps 1;

QY 1 KRDPQREYEDCRRCRCOEPRROHOCRLCRROQHGRGGMNPNORGS-GRYEEGE 59
 DB 215 RRPDSHREFSHRDRDEFYHSGSKRSRSEPRERDGRSGRQASPKRGSGPGREGS 274

QY 60 EQQ 62
 DB 275 EER 277

RESULT 15
 ID 081254 PRELIMINARY; PRT: 236 AA.
 AC 081254;
 DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE GLOBULIN-1 (FRAGMENT).
 OS Zea mays subsp. parviglumis.
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; Liliopsida; Poales; Poaceae; Zea.
 OX NCBI_Taxid=76912;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hilton H., Gaut B.S.,
 RT "Speciation and domestication in maize and its wild relatives:
 RT evidence from the Globulin-1 gene.";
 RL Genetics 0:0-0(1998).
 DR EMBL: AF064222; AAC31465.1; -;
 DR HSSP: PS0477; ICAU
 DR MENDEL; 31896; Zeama; 1188; 31896.
 DR INTERPRO: IPR000901; -;
 DR INTERPRO: IPR001113; -;
 DR PFAM: PF00546; Seedstore_7s; 1.
 DR PROSITE: PS00867; CPSASE_2; UNKNOWN_1.
 FT NON_TER 236
 SQ SEQUENCE 236 AA; 27050 MW; 36A7D94C5213F206 CRC64;

Query Match 22.4%; Score 79.5; DB 10; Length 236;
 Best Local Similarity 33.3%; Pred. No. 0.034;
 Matches 17; Conservative 8; Mismatches 21; Indels 5; Gaps 1;

QY 12 CRRKCEOEPRROHOCRLCRROQHGRGGMNPNORGSGRYEEGEEO 62
 DB 40 CVRCCEDRWQHPRCTLEOCREEFRSRHHEAD-----RSGEGSSFEDERRRK 85

Search completed: March 1, 2001, 16:09:07
 Job time: 1551 sec